SHORTER ARTICLES AND DISCUSSION

A SIMPLE TEST OF THE GOODNESS OF FIT OF MENDELIAN RATIOS

In actual experimentation the so-called Mendelian ratios, $3:1$, $9:3:3:1$, $9:3:4$, $9:7$, $15:1$, $27:9:9:9:3:3:3:1$, etc., are never exactly realized because of the errors of sampling inherent in all statistical work. Notwithstanding this fact, the best theoretical formulæ must be selected on the basis of these misleading experimental results.

Now the test of the validity of any Mendelian formula is twofold: the number of individuals found should agree with the number expected within the limits of experimental error,[1] the assumed germinal composition of the several groups of individuals should be capable of substantiation from a study of the soma of their offspring.

For the most part, Mendelians have been satisfied to judge the goodness of fit of the theoretical frequency to the empirical by inspection merely. More recently, however, attempts have been made to apply scientific tests to this problem. The first was that of Weldon,[2] but Professor Johannsen doubtless deserves the credit of having interested the few Mendelian workers who have taken the pains to calculate probable errors in this indispensable part of their work.

The test used by Professor Weldon and recommended in a much extended form by Professor Johannsen[3] is essentially the determination of the probable error of the number of individuals in one of the subgroups by the formula

$$S.D. = \sqrt{n \times p \times q},$$

where p is the chance of occurrence of an individual of any class, $q = 1 - p$, and n is the number of individuals. Thus the

---

[1] In some cases, valid reasons for discrepancy between calculated and observed frequencies may be shown. These factors should then be taken into account in calculating the theoretical numbers.

[2] Weldon, W. F. R., ''Mendel's Laws of Alternative Inheritance in Peas,'' *Biometrika,* 1: 228–254, 1902, especially pp. 233–234.

[3] Johannsen, W., ''Elemente der Exakten Erblichkeitslehre,'' pp. 402–410, 1909.

"probable error" of the number of individuals of any class, say p, is

$$Ep = .67449 \sqrt{npq}.$$

Now while Professor Weldon's use of this formula for the simple $3:1$ ratios seems quite proper, the same can not be said for Professor Johannsen's generalization. This is true for three reasons:

(*a*) The formula is valid only when neither n, p nor q is small. In polyhybrid ratios p or q may be relatively small.[4] It is then quite idle to use the probable error suggested, unless n be large, which unfortunately is generally not the case.

(*b*) Even when p is not so low as to render the use of the conventional formula for the probable error open to question, it is very laborious to calculate the probable errors for the frequency of each class.[5]

(*c*) It is not only cumbersome and laborious, but theoretically unjustified to test the validity of a given ratio by the determination of the probable error of one or of all of its individual component groups. The random deviations of the class frequencies are not independent, but correlated. We must have a usable criterion of the goodness of fit of the theory to the data as a whole.

Such a criterion was furnished several years ago by Pearson.[6] Its applicability to the problem of testing the goodness of fit of Mendelian ratios seems obvious, but since, as far as I can ascertain, it has nowhere been applied to this problem, it seems worth while to call the attention of students of genetics to its usefulness.

$$\chi^2 = S\{(o - c)^2/c\},$$

where o is observed frequency of any class, c is calculated frequency on the basis of Mendelian theory and S indicates a summation for the several classes distinguishable in the ratio under consideration.

P, a measure on the scale of 0 to 1 of the probability that

[4] For example, Johannsen (*loc. cit.*, p. 405) tables values for $p = 3/4$, $q = 1/4$ to $p = 63/64$, $q = 1/64$.

[5] See, for instance, the example given by Johannsen, *loc. cit.*, p. 396.

[6] Pearson, K., "On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling," *Phil. Mag.*, 50: 157–175, 1900.

the deviations from the theoretical frequencies may be reasonably supposed to be due to the errors of sampling, may be calculated from χ^2 by formulæ which need not concern us here, since its values for systems of frequency of 3–30 classes have been tabled.[7]   Hence the Mendelian has only the simple task of calculating χ^2 and looking up the value of P in Elderton's tables.

Illustrations will make method of computation and usefulness most clear.

ILLUSTRATION I.   DOUBLENESS AND PLASTID COLOR IN STOCKS
Saunders, *Journ. Gen.*, 1: 349–350, 1911

	Obs.	Calc.	$o-c$	$(o-c)^2$	$(o-c)^2/c$
Singles, White..........	1,666	1,615	51	2,601	1.611
Doubles, White..........	773	807	−34	1,156	1.433
Singles, Cream.........	790	807	−17	289	.358
	3,229	3,229	0		$\chi^2 = 3.402$

Whence, from the tables in *Biometrika*, and by interpolation,

$$n' = 3, \quad \chi^2 = 3, \quad P = .223130$$
$$n' = 3, \quad \chi^2 = 4, \quad P = .135335$$
$$\text{Diff.} = .087795$$
$$P = .223130 - .087795 \times .402 = .1878.$$

Thus only in about one case in five would the errors of sampling lead to divergences from theory as bad as this.   The theory is, as far as this evidence goes, *possible*, but certainly not demonstrated.

ILLUSTRATION II.   SEED FORM AND COLOR IN *Pisum*
Bateson and Killby, Report Evol. Com., 2: 77, 1905

	Obs.	Calc.[8]	$o-c$	$(o-c)^2/c$
Round, Yellow..................	4,926	4,883	+43	.3787
Wrinkled, Yellow...............	1,656	1,628	+28	.4816
Round, Green..................	1,621	1,628	− 7	.0301
Wrinkled, Green...............	478	542	−64	7.5572

$\chi^2 = 8.4476$, $P = .0384$.   Thus taking the data as they stand, it is impossible to regard the 9 : 3 : 3 : 1 ratio as satisfactorily

[7] Pearson, *loc. cit.*, gives a small table.   A much more comprehensive one is given by W. Palin Elderton, ''Tables for Testing the Goodness of Fit of Theory to Observation,'' *Biometrika*, 1: 155–163, 1901.

[8] These are not the calculated frequencies given by Bateson and Killby, but have been recalculated as closely as possible on the 9 : 3 : 3 : 1 ratio. Theirs are nine seeds short.

describing the facts.   But the great factor in the magnitude of χ^2 is the deficiency in the wrinkled green seeds, and the authors have suggested a reasonable biological explanation for this deficiency.

ILLUSTRATION III.   COLOR IN OATS
Nillson-Ehle, *fide* Baur. *Einf. Exp. Vererbungsl.*, pp. 66–67

	Obs.	Calc.	$o-c$	$(o-c)^2/c$
Schwarzspelzig..................	418	420	−2	.0095
Grauspelzig.....................	106	105	+1	.0095
Weissspelzig....................	36	35	+1	.0286

Thus $\chi^2 = .0476$ only. P is not tabled for $\chi^2 < 1$, since the probabilities of such deviations being due simply to errors of sampling are so enormously high.   Theory and observation could hardly agree more perfectly.

ILLUSTRATION IV.   BODY COLOR IN *Drosophila*
Morgan, *Journ. Exp. Zool.*, 13: 35, 1912

	Obs.	Calc.	$o-c$	$(o-c)^2/c$
Gray ♀	525	529	− 4	.030
Gray ♂.........................	340	265	+75	21.226
Yellow ♂.......................	194	265	−71	19.023
			$\chi^2 = 40.279$	

Here χ^2 is over 40, the odds against the deviations, being due to errors of sampling, are so enormously great that it is idle to express them in figures.   In short, the facts do not substantiate the hypothesis, and Professor Morgan has himself suggested possible reasons for the disagreement.

ILLUSTRATION V.   PARTIAL GAMETIC COUPLING IN SWEET PEAS
Bateson, Saunders and Punnett, Rep. Evol. Com., 4: 11

	Observed Number of Cases	Calculated on 7:1:1:7 Basis	Calculated on 7:1:1:7 Basis
Purple, long.....................	493	471	490
Purple, round....................	25	40	20
Red, long........................	25	40	20
Red, round.......................	138	130	151

For the $7:1:1:7$ basis, $\chi^2 = 12.7699$, $P = .0053$.   For the $15:1:1:15$ hypothesis, $\chi^2 = 3.6375$, $P = .3086$.   Thus the

chances are about $995:5$ or $199:1$ against the validity of the first hypothesis and only $69:31$, or about $2:1$, against the second.

ILLUSTRATION VI.   COLOR INHERITANCE IN *Antirrhinum*

Wheldale, Marryat and Sollas, Rep. Evol. Com., 15: 15

	Obs.	Calc.	$o-c$	$(o-c)^2/c$
I. Magenta.....................	399	361	+38	4.000
I. Magenta delila...............	122	120	+ 2	.033
I. crimson....................	121	120	+ 1	.008
I. crimson delila...............	38	40	− 2	.100
T. ivory......................	88	120	−32	9.075
T. ivory delila.................	35	40	− 5	.625
T. yellow.....................	33	40	− 7	1.255
T. yellow delila................	19	14	+ 5	1.786
	855	855	$\chi^2 = 16.852$	

Hence, $P = .0185$ or the chances are about $980:20$ against such discrepancies being chance deviations from the theory. Thus either the theory must be discarded or reasons for the discrepancies found.

A conspicuous advantage of this method of Pearson is that in its application the deviation of observation from theory for each class and the amount which this discrepancy contributes to χ^2 are under the worker's eye.

If used with the caution that should be exercised in the drawing of any conclusion from probable errors,[9] I believe that this well-known criterion of goodness of fit will prove most useful to Mendelians.

J. ARTHUR HARRIS

[9] Some biologists apparently seem to feel that the calculation of a statistical ''probable error'' covers all the biological sins which may be committed in the collection or manipulation of their data.